Police drowsy driving: predicting fatigue-related performance decay

Stephen M. James
Department of Criminal Justice and Criminology, Washington State University – Spokane, Spokane, Washington, USA, and
Bryan Vila
Department of Criminal Justice and Criminology, and Sleep and Performance Research Center, Washington State University, Spokane, Washington, USA

Abstract

Purpose – Fatigue associated with shift work is a well-established and pervasive problem in policing that affects officer performance, safety, and health. It is critical to understand the extent to which fatigue degrades officer driving performance. Drowsy driving among post-shift workers is a well-established risk factor yet no data are available about officer injuries and deaths due to drowsy driving. The purpose of this paper is to assess the impact of fatigue associated with work shift and prior sleep on officers’ non-operational driving using laboratory experiments to assess post-shift drowsy driving risks and the ability of a well-validated vigilance and reaction-time task to assess these risks.

Design/methodology/approach – Experienced police patrol officer volunteers (n = 78) from all four shifts of a medium-sized city’s police department were tested using a within- and between-subjects design to assess the impact of fatigue on individual officers, as well as the impact of different work shifts, on post-shift driving performance. Controlled laboratory experiments were conducted during which participants drove high-fidelity driving training simulators on two occasions: immediately following five consecutive 10:40-hour patrol shifts (fatigued condition) and again 72 hours after completing the last shift in a work cycle (rested condition).

Findings – Generalized linear mixed-model analyses of driving performance showed that officers working night shifts had significantly greater lane deviation during post-shift, non-operational driving than those working day shifts (F = 4.40, df = 1, 150, p = 0.038). The same method also showed that easy to measure psychomotor vigilance test scores for reaction time predicted both lane deviation (F = 31.48, df = 1, 151, p < 0.001) and collisions (F = 14.10, df = 1, 151, p < 0.001) during the simulated drives.

Research limitations/implications – Simulated driving tasks done by participants were generally less challenging than patrol or off-duty driving and likely underestimate the impact of fatigue on police driving post-shift or during extended shifts.

Originality/value – This is the first experimental research to assess the impact of shiftwork, fatigue, and extended shifts on police post-shift drowsy driving, a known risk factor for shift workers in general.

Keywords Fatigue, Police, Policy, Driving, Drowsy, Shift work

In total, 14 percent of officers in the first study of police fatigue (Vila, 2000) reported that they were always or usually tired at the beginning of their work shifts. This research focused on the other end of the work shift when officers tend to be even more tired, assessing the impact of work-related fatigue on an officer’s driving performance after a long work shift. Measuring fatigue at shift end provides an important performance indicator because that is when officers tend to either continue working on overtime status or drive home or to another destination. In any of these circumstances,
increasing levels of fatigue associated with time-on-task and time-awake effects may be expected to increase collision risks for officers and, consequently, others on the road with them. Although officer driving deaths and injuries that occur on the way home tend not to be counted as being “in the line of duty,” their consequences are equally tragic.

Drowsiness and falling asleep while driving are the result of fundamental human performance limits that affect us all. Drowsy driving is an important contributing factor in 22-27 percent of all traffic collisions in the USA (Pack et al., 1995; Klauer et al., 2006) and Great Britain (Parsons, 1986). Drowsy drivers tend to take longer to react, be less attentive to their environment, and have impaired decision-making skills (Jackson et al., 2013). Drivers who usually sleep \( \leq 5 \) hours per 24-hour period or snore (a symptom of sleep apnea) are significantly more likely to report driving drowsy during the previous 30 days (Wheaton et al., 2014). Furthermore, sleep-related collisions tend to happen more often at night or in the midafternoon when drivers are more likely to be sleepy (Pack et al., 1995; May and Baldwin, 2009). As a consequence, excessively sleepy individuals tend to have more – and more serious – traffic collisions (Tregear et al., 2009) and have much higher traffic fatality rates (National Transportation Safety Board (NTSB), 1990).

Thus, it is not surprising that a relatively large proportion of these collisions occur among workers driving home after night shifts (NTSB, 1990). The link between shift work and accidents has been established by evidence from many studies using different methods and types of populations (e.g. see Folkard et al., 2005; Akerstedt et al., 1994). A classic study of a large engineering company showed that working a night shift increased traffic collisions by 50 percent (Smith et al., 1994). Research on medical professionals showed that 79 percent of nurses working night shifts reported driving drowsy – nearly four times as great as was reported by day-shift nurses (Scott et al., 2007). Similarly, medical residents who had frequent on-call schedules had 6.7 times greater risk of motor vehicle collisions than those on less demanding schedules (Kowalenko et al., 2000; May and Baldwin, 2009). Historically, roughly 16 percent of police line-of-duty injuries each year – and 32 percent of fatalities – occurred in-vehicle collisions (Houser et al., 2004). In 2013, 36 percent of police officer deaths occurred in on-duty vehicle collisions (Federal Bureau of Investigation, 2015). To our knowledge, no data on officer deaths during the drive home from work has been collected or analyzed.

Among the general public, 41 percent report having fallen asleep or nodding off while driving, 4 percent within the past month (Wheaton et al., 2014), and 7 percent within the past six months (Tefft, 2010). Among US and Canadian police officers surveyed in 2004 by the AAA Foundation for Traffic Safety, 89 percent said drowsy driving is as dangerous as drunk driving and more than 93 percent said that it was a serious problem for both passenger car drivers and commercial drivers. In total, 95 percent believed that drivers who cause a crash because they are fatigued should be charged with a driving violation (AAA Foundation for Traffic Safety, 2004).

Despite this apparent awareness of the hazards associated with drowsy driving, nearly 50 percent of all police officers report having fallen asleep while driving, and about 25 percent report that this happens one to two times per month (Rajaratnam et al., 2011). This is not particularly surprising given that more than 40 percent of police officers report symptoms consistent with at least one sleep disorder and 29 percent have been assessed as excessively sleepy (Vila, 2006; Rajaratnam et al., 2011).

The research reported here had two goals: first, to assess the impact of operational fatigue associated with shift work and long work hours on police-officer driving safety, and second, to identify fatigue indicators that can be used to warn officers that their
driving performance is becoming degraded by fatigue – preferably well before they are impaired. In order to accomplish these goals, we compared the driving performance of 78 experienced police officers during a 30-minute drive in a high-fidelity driving training simulator that is widely used by police (Patrol Sim IV, L3 Corporation) during two conditions: while fatigued after working five consecutive 10:40-hour shifts in the field, and at the same time of day after three consecutive days off. This driving task came at the end of a five-hour experimental session designed to mimic an officer’s work activities. This end-of-study driving task was designed to address the risks of driving home at the end of a long work week, or driving during five hours of overtime following a 10:40-hour work shift.

Sleep, shift work, fatigue, and driving
Sleep, like food, water, and oxygen, is a fundamental biological need for every living organism that has even a rudimentary brain. Without sufficient sleep the signal quality between groups of brain cells declines until they hit a limit where they fall asleep – sometimes for a fraction of a second, sometimes for several seconds or more (Boyle et al., 2008). From a practical standpoint, this means that regions of the brain that have been working the hardest tend to go offline sooner, and that no amount of willpower or determination will keep them from doing so. Drivers feel increasing drowsiness when their brain cells begin going offline. However, by the time they begin struggling with drowsiness as a consequence of sleep-related fatigue, they are already substantially impaired (Belenky et al., 2003; Van Dongen et al., 2003).

Sleep-related performance impairment due to fatigue is a function of three primary causal factors: time of day, time awake without sleep, and the amount of sleep obtained during recent days (commonly referred to as “sleep debt”). Time of day is critical because, despite the 24/7 demands of contemporary life, humans’ inherent circadian rhythms push us to sleep at night and stay awake during the day. That is why people who have been up all night tend to feel more alert as the sun comes up than they did at 2 a.m. Time awake has a cumulative effect on how fatigued or drowsy we feel as groups of our brain cells become increasingly sleep deprived and begin falling asleep. Prior sleep affects how tired we are at the beginning of our time awake. Each consecutive day with insufficient sleep adds to our “sleep debt,” lowering the threshold at which fatigue will begin affecting performance. This is akin to the decline in available credit on a charge card that occurs as holiday purchases outpace payments. Although the impact of each of these three causal factors on fatigue is reasonably straightforward, their systematic interactions are very complex (Hursh et al., 2004; Van Dongen et al., 2007; Van Dongen and Dinges, 2005).

It is challenging to predict how fatigued a person is at a particular point in time because it requires knowing the current time of day, how long they have been awake, and how much sleep they have had during the recent past. And the magnitude of each of these factors is constantly shifting as well because of the circadian rhythms that coordinate the functions of the body’s complex biological processes across each day. As we move into the late evening hours when our bodies are programmed to fall asleep, the impact of time awake and prior sleep on fatigue becomes increasingly strong. Later, as we move past daybreak, the impact tends to become weaker.

On top of this ever-shifting system of interactions, the manner in which fatigue affects driving performance shifts from moment to moment. As a driver becomes more drowsy, the probability of lapses of attention caused by clusters of brain cells falling asleep tends to increase, but the timing and the duration of those attentional lapses are
random. This means that drowsy drivers’ attention to the roadway around them tends to drift in and out. As a result of this unpredictable waxing and waning of attention, people often avoid the consequences of driving drowsy simply because their lapses of attention do not co-occur with a potential hazard. For example, if you experience a momentary attentional lapse while driving on a road and regain alertness before the car runs off the road or collides with something, you may not even realize you dozed off. But you collide if, for example, the lapse happens just as someone pulls in front of you, the road turns away from your current vector, or a vehicle brakes unexpectedly in front of you.

Although lapses of attention are often the primary connection between fatigue and driving performance, fatigue also can increase the probability of accidents, injuries, and errors on the job by interfering with mood, cognitive ability, and risk-taking propensity (see reviews by Durmer and Dinges, 2005; Folkard and Tucker, 2003). For example, when one considers the proximate causes of the 7,117 officer-involved injury and fatal collisions in California from 1997 to 2007, the three leading “primary causal factors” were unsafe speed (35 percent), automobile right of way (12 percent), and improper turning (10 percent) (California Commission on Peace Officer Standards and Training, 2009, p. 12). However, sleep-related performance factors upstream from these proximate causes include mood (which can bias decision making), cognitive ability (which interferes with perception, decision making, and task execution), and risk-taking propensity (which affects the behavioral choices that officers make).

Data and methods
Sample size and power analysis
Our sample size was selected in a way that placed a premium on finding a relationship between work-related fatigue and performance in critical operational tasks, such as driving, if it existed. This was appropriate because the consequences of poor performance are potentially catastrophic and the potential advantages of being able to better manage fatigue are great. Our challenge was to accomplish this in spite of concerns that variation associated with making the simulations appropriately realistic might mask performance effects. An earlier pilot study using the driving and deadly force simulators in a similar research design with 27 subjects working the graves shift [1] established that they could detect highly significant effects for lane deviation and vigilance (Waggoner et al., 2012, pp. 1576-1577). The graves shift tends to have a greater impact on worker performance than other work shifts of similar length because it is contrary to both the body’s natural tendency to sleep at night and be awake during the day, and the rhythms of social life among the majority of the population. We addressed this problem by accompanying the driving simulations with high-stimulus-load psychomotor vigilance tests (PVT) which provide very sensitive and well-validated assays of degradation in vigilance and reaction time (rt). They are also strong predictors of driving performance degradation. Thus, our approach to sample size was conservative—and we included the PVT as a reliable independent variable.

Our analysis employed mixed method models using time asleep measured by actigraphy as an independent, repeated measure which is an inverse measure of sleep loss. The performance scores in PVT, deadly force, cognitive battery, both driving tasks, and social interaction were multivariate dependent measures. We set statistical significance at a Type I error rate of $\alpha = 0.05$ (one-tailed). Power calculations revealed that 80 subjects divided into two groups would allow us to detect an effect of the independent variables with an effect size as small as 0.2 with more than 80 percent
power (Lipsey, 1990, p. 92). Thus, with 80 subjects, we were confident that our study was sufficiently sensitive to meet Cohen’s criterion of less than 0.2 for the smallest worthwhile effect size (Cohen, 1988).

Participants

All of the experienced police patrol officer volunteers (n = 80) were assigned to full-time patrol work in the Spokane, Washington, Police Department. SPD is a medium-large municipal police agency with about 250 sworn, full-time officers. Of the 80 officers participating in the study, 78 completed both conditions. Only their data are included in the analyses presented here. These volunteers represented about half of the patrol force assigned to each of four shifts. Participants were asked to keep a normal routine during the study, with the exception of not accepting overtime assignments during the 72 hours prior to the rest condition. Participants’ sleep was tracked objectively by wrist actigraphy (using the SBv2 ReadiBand by Fatigue Science, Vancouver, BC). These watch-sized devices record movement in one-minute bins as either sleep or awake. Subjective self-reports of sleep were also collected using sleep diaries for five days prior to each condition.

Roughly half of the officers worked night shifts that required them to get all or most of their sleep during the daytime, when it is most difficult for humans to fall asleep, sustain sleep, or sleep soundly (see reviews by Folkard et al., 2005; Folkard and Tucker, 2003). As Table I shows, similar numbers of officers were recruited from all four patrol work shifts. On average, officers in the study were 40.4 years old and had 14.5 years of experience. Those on the days (06:00-16:40) and power (10:00-20:40) shifts (a.k.a. “night sleepers”) tended to be nine years older and have eight years more experience than those on the Swing (16:00-02:40) and Graves (20:00-06:40) shifts (a.k.a. “day sleepers”). Slightly fewer than one out of eight officers in the study were women, and all but one worked on either the power or swing shift. On average, officers’ one-way commute from home to work took about 24 minutes.

Implications of non-random sampling

Representativeness of the participants is a potential issue because of the non-random sampling design with regard to both the agency from which officers were recruited and the generalizability of results to police officers in the USA. Officers were recruited directly via posters and peer-led discussions during roll call with the permission of the administration but without any encouragement from supervisors or superiors. During

<table>
<thead>
<tr>
<th>Work shift</th>
<th>Days (06:00-16:40)</th>
<th>Power (10:00-20:40)</th>
<th>Swing (16:00-02:40)</th>
<th>Graves (20:00-06:40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>n = 17</td>
<td>n = 21</td>
<td>n = 17</td>
<td>n = 23</td>
</tr>
<tr>
<td>Fatigued condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rested condition</td>
<td>n = 18</td>
<td>n = 22</td>
<td>n = 17</td>
<td>n = 21</td>
</tr>
<tr>
<td>Male</td>
<td>100%</td>
<td>81.40%</td>
<td>76.50%</td>
<td>95.50%</td>
</tr>
<tr>
<td>Age</td>
<td>Mean 46, SD 7.4</td>
<td>Mean 43.9, SD 7.5</td>
<td>Mean 35.4, SD 5.7</td>
<td>Mean 36.3, SD 5.8</td>
</tr>
<tr>
<td>Years sworn service</td>
<td>21, SD 6.4</td>
<td>16.6, SD 5.2</td>
<td>10.2, SD 6.5</td>
<td>10.7, SD 6.1</td>
</tr>
<tr>
<td>Commute one way (min)</td>
<td>00:25, SD 00:15</td>
<td>00:25, SD 00:13</td>
<td>00:23, SD 00:06</td>
<td>00:23, SD 00:11</td>
</tr>
</tbody>
</table>

Table I. Characteristics of police-officer research participants on each work shift.
testing, watch commanders agreed to ensure that participants were relieved in time to come to the laboratory immediately after their last work shift in a cycle.

Representativeness of subjects within the department
We believe that our quasi-experimental research design and sample size made it very likely that the sample is representative of the department’s patrol force as a whole. First, roughly half of the patrol force volunteered for the study and the within-subjects analysis of the impact of fatigue on performance meant that each subject served as his or her own control. Second, as is described in more detail below, officers were tested experimentally in two conditions: immediately after the last shift in a five-day sequence of 10:40-hour shifts (fatigued condition), and at the same time of day several weeks earlier/later after three days off (rested condition). Random counterbalancing was used to avoid order effects. As a consequence of these two factors, the data collected comes from a highly controlled sample of 50 percent of the study population. The population was roughly evenly stratified between the four, non-overlapping patrol shifts, and each of those shifts contained a different age cohort with substantial differences in age and experience between night sleepers and day sleepers (see Table I). There was no significant difference in age nor years of experience between officers that took part in the study, the remaining 50 percent of the patrol division that did not participate, those not assigned to patrol, nor the agency as a whole. The sample of officers in this study was representative of the agency, as can be seen in Table II.

External generalizability to other agencies
All of the officers who participated in this study came from the same, medium-sized policing agency. This was necessary in order to control their exposure to the central independent variable in this study, fatigue obtained in a genuine police environment. All of our subjects worked the same length shifts in the same mixed urban environment and for the same agency. Within their strata of the sample, they also worked the same time of day. This raises the question of how generalizable our results are likely to be to other agencies in the USA.

In order to assess generalizability, we compared the study population with national-level data on both demographics in similarly sized policing agencies and workload as indicated by local crime rates and staffing levels. We also compared them with prior studies on officer fatigue, wellness, and operational performance in terms of demographic, experiential, and sleep characteristics. Because roughly 75 percent of all line-of-duty accidental deaths are vehicle related, we also compared participants’ age, sex, and years of experience with those of officers who were killed accidentally in the line of duty during the last year for which data were available. The comparisons made below establish that our results are likely to be generalizable, although the level of work-related fatigue associated with their patrol work may be somewhat higher due to the agency’s substantially lower-than-average staffing ratios.

<table>
<thead>
<tr>
<th>Table II.</th>
<th>Participants Mean</th>
<th>Participants SD</th>
<th>All patrol Mean</th>
<th>All patrol SD</th>
<th>Non-patrol Mean</th>
<th>Non-patrol SD</th>
<th>All officers Mean</th>
<th>All officers SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>40.4</td>
<td>8.0</td>
<td>40.3</td>
<td>7.3</td>
<td>43.5</td>
<td>7.5</td>
<td>41.8</td>
<td>7.5</td>
</tr>
<tr>
<td>Years sworn service</td>
<td>14.5</td>
<td>7.5</td>
<td>13.0</td>
<td>7.2</td>
<td>17.0</td>
<td>7.2</td>
<td>14.8</td>
<td>7.5</td>
</tr>
</tbody>
</table>
Few data are available at the national level for comparing our participants’ individual-level variables of interest, but there are sufficient indicators of workload. As Table III shows, when compared with available national data, the proportion of male subjects is within 1.0 percent of the national average and the community rate for violent UCR part I crimes is within 2.5 percent of the national average for cities with populations $\geq 100,000$. Part I property crimes, however, are 2.3 times greater than the national average for cities $\geq 100,000$, and the sworn officer staffing level per 1,000 population is 64.3 percent lower than the national average[2]. These coarse comparisons suggest that patrol officers in this study tend to have relatively high workloads when compared with urban police agencies. Higher workload may be expected to increase on-duty fatigue levels, which could result in officers being more fatigued during the on-duty (i.e. fatigued) condition. This increases the likelihood of capturing fatigue-related effects on driving performance if they exist.

Officers participating in the current study were also very similar to those from each of the prior studies of officer fatigue, wellness, and performance. Table IV compares research participant similarity in age, years of sworn service, commute time to work, and self-reported recent average sleep per day from the following studies:

- the highly detailed ongoing longitudinal BCOPS study (Violanti et al., 2009; Charles et al., 2007) collects data on hundreds of biological, psychological, work hours, and demographic variables from nearly all of the officers in a single medium-large police department (patrol and other assignments);
- the Harvard police sleep study (Rajaratnam et al., 2011) used clinical assessments of officers from two agencies as well as survey studies of thousands of officers from across the USA (patrol and other assignments, plus 3 percent of subjects from Canada and Australia);
- the NIJ police work-shift study (Amendola et al., 2011) used experimental and clinical measures with two larger-sized police agencies (patrol officers only); and
- the NIJ tired-cops study (Vila, 2000; Vila et al., 2002) used experimental and clinical measures with four medium-sized police agencies in different regions of the USA (patrol officers only).

As Table V shows, the participants’ demographic and experiential characteristics also are very similar to those from officers who died accidentally in 2013, the last year for

<table>
<thead>
<tr>
<th>Officer/Agency/Community characteristics</th>
<th>Study participants</th>
<th>National comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>88.5%</td>
<td>87.5%</td>
</tr>
<tr>
<td>Agency staffing per 1,000 population</td>
<td>1.4</td>
<td>2.3</td>
</tr>
<tr>
<td>2012 crime rates compared to all US cities $\geq 100,000$ population $^a$</td>
<td>Violent 645</td>
<td>Violent 662</td>
</tr>
<tr>
<td></td>
<td>Property 8,730</td>
<td>Property 3,718</td>
</tr>
</tbody>
</table>

Table IV.

Representativeness: comparison of officer sleep characteristics and prior research on police sleep and fatigue.

<table>
<thead>
<tr>
<th>Source</th>
<th>Current study</th>
<th>Violanti BCOPS 2012</th>
<th>Rajaratnam 2011&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Amendola work shift 2011</th>
<th>Vila tired cops 1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>78</td>
<td>419</td>
<td>4,957</td>
<td>275</td>
<td>1998</td>
</tr>
<tr>
<td>Population characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Age</td>
<td>40.2</td>
<td>7.9</td>
<td>42.1</td>
<td>7.3</td>
<td>38.5</td>
</tr>
<tr>
<td>Years sworn service</td>
<td>14.5</td>
<td>7.4</td>
<td>15.3</td>
<td>7.2</td>
<td>12.7</td>
</tr>
<tr>
<td>Commute time (min)</td>
<td>24.0</td>
<td>11.9</td>
<td>16.4</td>
<td>5.4</td>
<td>na</td>
</tr>
<tr>
<td>Recent avg. sleep/day (hr)</td>
<td>6.5</td>
<td>1.3</td>
<td>6.1</td>
<td>1.2</td>
<td>na</td>
</tr>
</tbody>
</table>

Note: <sup>a</sup>Commute and sleep data were unavailable from the authors at the time this paper was written.
which LEOKA data are available (FBI, 2015). Historically, roughly half of all on-duty officer deaths in the USA are accidental, and vehicle-related deaths account for about 75 percent of all accidental deaths (Houser et al., 2004).

Research protocol and measurement
This study was part of a larger research project assessing the impact of work-shift related fatigue on officers’ performance of simulated critical operational tasks: deadly force judgment and decision making, driving, cognition, and “non-operational” driving. (Recall that so-called non-operational driving is generally easier than patrol driving and is therefore a conservative measure of the impact on operational driving during extended shifts.) The topic of this paper, non-operational driving (a.k.a., the drive home), was the last experimental task performed (see Table IX) during each five-hour-long experiment. Fatigue was obtained during real police patrol work, but all experimentation was conducted under laboratory conditions (controlled for access, consistent light, sound, and temperature) in the Simulated Hazardous Operational Tasks laboratory in Washington State University’s Sleep and Performance Research Center.

During the non-operational drive that began about 4.3 hours into the research protocol, participants drove a simulated Toyota Camry in a fixed-base, high-fidelity driving training simulator (L3 Communications, MPRI Patrol Sim IV, see Plate 1). The simulator was adapted for driving measurement purposes by installing additional software and hardware (Moore et al., 2009). The simulator itself used both software and hardware to realistically replicate the mechanics and driving characteristics of an actual car. Because changes in lane position are important for measuring driving performance, we previously assessed the realism of the manner in which changes in steering-wheel position are translated into changes in-vehicle heading, finding that they are realistically complex and reliable. (see Forsman et al., 2013, for a complete description.)

Officers drove standardized daytime driving scenarios in the simulator on a rural highway with no adverse weather or other vehicles. The 28-mile course was fixed, with ten 0.5-mile straight and uneventful segments, and eight 0.5-mile curved and uneventful road segments (see Plate 1). Random events were placed at five to seven other locations around the track in which pedestrians or dogs would cross the road. Participants were instructed to maintain a speed of 55 mph around the track, stay in the same lane, and brake without swerving if a pedestrian or dog walked into the roadway at a leisurely pace. At that speed, an alert person could readily stop before striking them, so this was a conservative challenge. Vehicle data (e.g. speed, acceleration, braking, steering-wheel angle, heading, etc.) were sampled at 72 Hz.

Participants drove on two separate occasions, once following the last of five consecutive 10:40-hour patrol shifts in their usual work week (fatigued condition) and again several weeks later at the same time of day 72 hours after completing the last shift in a work week (rested condition). The minimum duration between test conditions was 12 days, thus allowing a full uninterrupted duty/rest cycle to take place between testing.

<table>
<thead>
<tr>
<th>Current study</th>
<th>LEOKA 2013 accidental deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>40.2</td>
</tr>
<tr>
<td>Years sworn service</td>
<td>15</td>
</tr>
<tr>
<td>% Male</td>
<td>89</td>
</tr>
</tbody>
</table>
This ensured that the added workload from the first testing condition did not impact the subsequent testing condition. Subjects’ availability for the experiments varied somewhat due to work and personal schedules. As a consequence, the mean number of days between conditions was 32.4 (SD = 24.5).

In each condition, officers drove identical courses in which only the location and timing of the pedestrian or dog events were varied along straight sections of the roadway. Conditions were counterbalanced to control for order effects between the two conditions by randomly assigning participants to have either the fatigued (51 percent) or rested (49 percent) condition first.

Each drive was preceded and followed by a ten-minute PVT (Pulsar model 2.0.5.9, Philadelphia, PA). The PVT is a well-validated and simple reaction-time task with high-stimulus density. It measures participants’ ability to sustain attention (Lim and Dinges, 2008) and is the most widely used objective research measure of sleep-related fatigue. During each PVT, officers were asked to watch for a number to appear on a computer screen, then press the keyboard spacebar as quickly as possible. The number appeared at random intervals of from two to 11 seconds. Once it appeared, the number on screen counted up in milliseconds until the spacebar was pressed.
Participants completed from 72 to 98 tests during each experiment (Mean = 93.8, SD = 2.7). As is generally done in sleep research, a RT of less than 200 ms was considered a “false start” and a RT longer than 500 ms was considered a lapse in attention. We also administered the Karolinska Sleepiness Scale (KSS), a brief, well-validated self-report assessment of subjective sleepiness on a scale from 1 (extremely alert) to 9 (extremely sleepy – fighting sleep) which also is widely used in sleep research (Kaida et al., 2006).

Results
Data analysis
IBM SPSS (v. 22.0.0.0, New York, NY) was used for statistical analysis. Since the research protocol required multiple observations per participant, the data potentially violated the assumption of independence. To allow for this, a generalized linear mixed model with a hierarchal structure was used. These types of models use a combination of ANOVA and regression and are designed to deal with data that have multiple layers or that are hierarchical. In this case, the “layers” of data are participants and individual responses. Thus, individual data points from within each experimental group cannot be considered independent from each other because they may be grouped or “clustered” around the responding participant.

Officer sleep
Across both conditions (rested and fatigued) there was no significant difference between day shifts (days and power) and night shifts (swing and graves) in the amount of sleep officers received prior to participating in the study. Participants assigned to the days shift slept an average of 7.0 hours (SD = 1.1), power 7.4 hours (SD = 1.1), swing 7.2 hours (SD = 1.5), and graves shift 7.2 hours (SD = 1.7) in a 24-hour period.

There were highly significant differences in the amount of sleep officers received per 24-hour period in the 72 hours preceding their participation in the rested- and fatigued-condition experiments as calculated by one-way ANOVA. As Table VI details,

<table>
<thead>
<tr>
<th>All shifts</th>
<th>Days 06:00-16:40</th>
<th>Power 10:00-20:40</th>
<th>Swing 16:00-02:40</th>
<th>Graves 20:00-06:40</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Rested</td>
<td>40.2</td>
<td>7.9</td>
<td>42.1</td>
<td>7.3</td>
</tr>
<tr>
<td>Fatigued</td>
<td>6.50</td>
<td>1.03</td>
<td>6.24</td>
<td>0.83</td>
</tr>
<tr>
<td>Difference</td>
<td>-1.38***</td>
<td>-1.53***</td>
<td>-0.47</td>
<td>-1.33*</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rested</td>
<td>10.44</td>
<td>6.27</td>
<td>12.89</td>
<td>2.71</td>
<td>16.75</td>
<td>5.74</td>
<td>6.18</td>
<td>4.74</td>
<td>5.07</td>
<td>0.49</td>
</tr>
<tr>
<td>Fatigued</td>
<td>18.38</td>
<td>4.23</td>
<td>15.86</td>
<td>2.64</td>
<td>18.38</td>
<td>5.45</td>
<td>18.76</td>
<td>3.74</td>
<td>19.88</td>
<td>3.79</td>
</tr>
<tr>
<td>Difference</td>
<td>7.94***</td>
<td>2.97**</td>
<td>1.63</td>
<td>12.58***</td>
<td>14.81***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hours awake prior to testing</th>
<th>Rested</th>
<th>Fatigued</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rested</td>
<td>10.44</td>
<td>18.38</td>
<td>7.94***</td>
</tr>
<tr>
<td>Fatigued</td>
<td>20.35</td>
<td>10.38</td>
<td>10.02</td>
</tr>
</tbody>
</table>

Notes: *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001

Table VI.
on average, days-shift officers slept 1.6 hours more each day while off duty before the rested condition ($f = 28.9, df = 1, 33, p < 0.001$); swing shift officers slept 1.3 hours more each day while off duty before the rested condition ($f = 7.3, df = 1, 30, p = 0.011$); and graves-shift officers slept 2.2 hours more each day while off duty before the rested condition ($f = 36.35, df = 1, 38, p < 0.001$). Each of these groups slept significantly less during their work week, with graves-shift officers sleeping the least and relying more on their off-duty days to catch up on sleep.

However, officers on power shift did not significantly vary the amount of sleep they acquired between on-duty and off-duty workdays. This likely is because their shift (10:00-20:40) allows ample time to sleep at night without the burden of getting up early enough for the 06:00 days shift. On average, they slept only 0.5 hours more each day while off duty before the rested condition than during on-duty days ($SD = 0.8, f = 2.1, df = 1, 40, p = 0.151$).

Officer age
The age of officers participating in this study ranged from 28 to 58 years old. As Table VI shows, although the age distribution appears normal across the sample, age is not normally distributed between work shifts. Officers at the upper age range tend to be on days and the younger officers tend to work nights. Age can interact with circadian disprution to affect performance. We tested this hypothesis using a general linear mixed-effects model to control for condition, time of day, prior sleep, and hours awake. Of the four shifts only the power shift did not report a significant relationship between age and rt during the PVT. The other three shifts reported a significant inverse relationship; however, the effect size of age on rt was small. Each additional year in age reduced day-shift officers’ average rt by 2.9 ms ($f = 7.11, df = 4, 29, p = 0.012$), power-shift officers by 2.4 ms per year ($f = 5.40, df = 5, 34, p = 0.026$), and graves shift by 6.8 ms per year ($f = 6.77, df = 5, 36, p = 0.013$). When examining the effect of age on driving performance measures, the only significant relationship was with the graves shift and lane deviation. This inverse relationship accounted for one centimeter of variation per year of age. Due to these small effect sizes, age was not considered further in these analyses.

Collisions by work shift
Generalized linear mixed-model analyses of driving performance showed that officers working night shifts (swing and power shifts) had significantly greater lane deviation during post-shift, non-operational driving than those working day shifts ($f = 4.40, df = 1, 150, p = 0.038$). More simulated collisions occurred in the fatigued condition compared to the rested condition, and this effect was much stronger for day-sleeping officers (days and power shifts). Among days-shift officers, 11 percent of rested-condition drives ended in a collision vs 19 percent of fatigued-condition drives. Officers assigned to the power shift had slightly fewer collisions during the fatigued condition (14 percent of fatigued-condition drives ended in a collision vs 15 percent of rested-condition drives). However, one of the power-shift officers had two collisions in one fatigued-condition drive, making the total number of collisions in the fatigued condition higher than those during the rested condition. The effect of fatigue during the swing shift was much greater than during either of the night-sleeping shifts. Among swing-shift officers, only 6 percent of rested-condition drives ended in a collision, compared with 41 percent of fatigued-condition drives. Among grave-shift officers, 14 percent of rested-condition drives ended in a collision, compared to 22 percent of
fatigued-condition drives. Interestingly, collisions increased at the same rate between conditions for both graves and days officers, although the graves-shift officers started at a higher base rate (14 percent collision rate for rested grave-shift officers vs. 11 percent collision rate for days-shift officers). Figure 1 illustrates these findings (Table VII).

Prior sleep, fatigue predictors, and driving performance

A within- and between-subjects design was used to compare participants’ performance by shift in the fatigued and rested conditions using well-established driving performance metrics, such as the standard deviation of a vehicle’s position within the lane, leaving the assigned lane, and braking latency that are strong predictors of the likelihood of a collision. The design also was used to assess the predictive power of the relatively simple PVT and KSS instruments. PVT and KSS scores are potentially important practical indicators of fatigue because predicting fatigue in real time is complex—one must assess the combined influence of time of day, time awake, and prior sleep. If PVT or KSS are strong predictors, they would be relatively easy to use in the field to help manage fatigue.

In the between-subjects analysis, there was considerable variation between subjects in PVT reaction times (rt), PVT lapses of attention (rt > 500 ms), KSS self-assessed sleepiness scores, and driving performance. A general linear mixed-effects model was used to control for between-subject variation by using repeated measures to compare

![Figure 1. Comparison of rested- vs fatigued-condition collisions by work shift/work hours](image)
participants’ fatigued performance against their own rested performance. In other words, the method allowed each participant to serve as his or her own control.

Prior sleep
While controlling for time-of-day effects, we tested a model containing the average amount of sleep a participant had for the previous 72 hours and how long he or she had been awake at testing. The model significantly predicted probability of a collision ($f = 4.00$, df = 3, 144, $p = 0.009$). Table VIII shows the breakdown of this relationship.

### Table VII.
Rest vs fatigued-condition relationships between participants’ regular work shift and work hours and performance variables measuring reaction time, vigilance, self-assessed alertness, driving precision, braking latency, and collisions

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>SD</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days (0600-16:40)</td>
<td>23.24</td>
<td>2.62</td>
<td>18.72</td>
<td>2.49</td>
<td>20.89</td>
<td>3.40</td>
<td>$f = 26.60$, df = 1, 32, $p &lt; 0.001^{***}$</td>
</tr>
<tr>
<td>Power (10:00-20:40)</td>
<td>22.88</td>
<td>3.67</td>
<td>21.32</td>
<td>2.49</td>
<td>22.12</td>
<td>3.21</td>
<td>$f = 2.53$, df = 1, 32, $p = 0.120$ (ns)</td>
</tr>
<tr>
<td>Swing (16:00-02:40)</td>
<td>23.48</td>
<td>4.52</td>
<td>19.50</td>
<td>3.77</td>
<td>21.49</td>
<td>4.56</td>
<td>$f = 7.30$, df = 1, 30, $p = 0.011^{**}$</td>
</tr>
<tr>
<td>Graves (20:00-06:40)</td>
<td>25.70</td>
<td>4.63</td>
<td>18.49</td>
<td>2.93</td>
<td>21.55</td>
<td>5.16</td>
<td>$f = 36.35$, df = 1, 38, $p &lt; 0.001^{***}$</td>
</tr>
</tbody>
</table>

Notes: $a^{KSS}$ scores (1 = extremely alert, 3 = alert, 5 = neither alert nor sleepy, 7 = sleepy – but no difficulty remaining awake, 9 = extremely sleepy – fighting sleep). $*p \leq 0.05; **p \leq 0.01; ***p \leq 0.001$
between the four work shifts, which are illustrated in Figure 2. As was observed when comparing collisions during rested vs fatigued conditions, the day, swing, and graves shifts showed significant between-condition differences in the amount of sleep officers obtained prior to the experimental protocol, but the power-shift effect was not significant.

**Fatigue predictors**

The utility of the objective PVT and the subjective KSS self-report for use as indicators of fatigue in police officers was assessed by comparing each instrument’s ability to predict fatigue-related impairment as calculated from direct measures of the three most important factors causing sleep-related fatigue. A generalized linear mixed model was used to analyze PVT results for the four work shifts, while controlling for time awake before testing, the amount of sleep participants had obtained during the previous 72 hours, and condition (rested or fatigued). We found that officers on the days shift had significantly faster PVT rt \((f = 2.87, \text{df} = 6, 140, p = 0.011)\) and fewer lapses of attention (rt > 500 ms) \((f = 2.71, \text{df} = 6, 140, p = 0.016)\) than those assigned to the graves shift.

KSS scores for the four work shifts were analyzed using the same method as for the PVT. Time awake before testing, the amount of sleep participants had obtained during the previous 72 hours, and work shift were all significant in the model predicting KSS.

![Figure 2. Average hours of sleep obtained by participants during the 72 hours preceding experiments in each condition](image)
scores \( (f = 12.6, \ df = 5, 144, \ p < 0.001) \). Officers’ work shifts had the greatest effect using this statistical model \( (f = 14.0, \ df = 3, 144, \ p < 0.001) \). Given that both PVT and KSS were indicators of fatigue levels, we next assessed their respective abilities to predict officer driving performance.

**PVT and KSS as driving performance predictors**

As was discussed previously, driving performance was assessed in the experiments using the standard deviation of lane position and braking latency, both of which are well-validated predictors of collisions. Assessing these leading indicators of increasing risk of collision makes it possible to prevent collisions by intervening before risks exceed some predetermined threshold. However, measuring lane deviation in the field is difficult because lane position requires the ability to measure position relative to constantly changing roadway edges and lane markers. Similarly, braking latency requires knowledge of the time a stimulus to provoke breaking first appeared (e.g. the brake lights of a vehicle ahead or a person or vehicle pulling onto the roadway ahead). Therefore, PVT and KSS would currently be easier measures to implement than either lane deviation or braking latency.

As expected, in our generalized linear mixed-model analyses greater lane deviation significantly predicted collisions \( (f = 25.06, \ df = 1, 150, \ p < 0.001) \). This effect was particularly strong for both the day shift \( (f = 23.62, \ df = 1, 32, \ p < 0.001) \) and the graves shift \( (f = 19.28, \ df = 1, 41, \ p < 0.001) \). Braking latency also significantly predicted collisions \( (f = 4.54, \ df = 1, 146, \ p = 0.035) \), with the strongest effect observed for the swing \( (f = 6.43, \ df = 1, 32, \ p = 0.016) \) and graves shifts \( (f = 7.12, \ df = 1, 39, \ p = 0.011) \). When considered together, both lane deviation \( (f = 34.24, \ df = 1, 145, \ p < 0.001) \) and braking latency \( (f = 13.65, \ df = 1, 145, \ p < 0.001) \) significantly predicted collisions \( (f = 19.58, \ df = 2, 145, \ p < 0.001) \). All but the power shift showed strong effects between lane deviation, braking latency, and collisions; day shift \( (f = 11.00, \ df = 2, 31, \ p < 0.001) \), swing shift \( (f = 5.62, \ df = 2, 21, \ p = 0.008) \), and once again the strongest effects were found in the graves shift \( (f = 21.33, \ df = 2, 38, \ p < 0.001) \).

Using the same method as was used above for lane deviation and braking latency, we found that changes in an officer’s PVT \( t \) \( (f = 14.10, \ df = 1, 151, \ p < 0.001) \), PVT lapses \( (f = 5.33, \ df = 1, 151, \ p = 0.022) \), and KSS scores \( (f = 4.80, \ df = 1, 150, \ p = 0.030) \) were all significant predictors of collision.

A generalized linear mixed model was then used to analyze the relationship between lane deviation and PVT results. PVT scores for \( t \) predicted both lane deviation \( (f = 31.48, \ df = 1, 151, \ p < 0.001) \) and collisions \( (f = 14.10, \ df = 1, 151, \ p < 0.001) \) during the simulated drives. Increased PVT \( t \) predicted participants’ increased lane deviation \( (f = 50.83, \ df = 1, 150, \ p < 0.001) \) and more lapses of attention also predicted increased lane deviation \( (f = 25.86, \ df = 1, 150, \ p < 0.001) \). This effect was strongest with the graves shift for both \( t \ (f = 31.33, \ df = 1, 41, \ p < 0.001) \) and lapses of attention \( (f = 25.86, \ df = 1, 41, \ p < 0.001) \). When controlling for shift, a generalized linear mixed model also showed that increased PVT \( t \) significantly predicted greater braking latency \( (f = 2.72, \ df = 4, 143, \ p = 0.032) \), as did increased PVT lapses of attention \( (f = 6.48, \ df = 1, 146, \ p = 0.12) \). As before, the graves shift produced the strongest relationships \( (f = 5.05, \ df = 1, 39, \ p = 0.030) \).

KSS also predicted lane deviation using these methods, with higher self-reported fatigue scores predicting greater lane deviation by participants \( (f = 4.10, \ df = 1, 149, \ p = 0.047) \).


Discussion
The drive home experiment assessed the likely impact of fatigue on driving immediately after the last shift in a five-day work cycle of 10:40-hour shifts, making it applicable to both drowsy driving during the commute home and driving during extended work shifts. Because direct prediction of fatigue is complex, we used three approaches to assessing participant’s fatigue impairment: first, lane deviation and braking latency in a high-fidelity police driving simulator; second, a well-validated instrument for objective measurement of vigilance and rt (PVT); and third, a subjective self-report of sleepiness (KSS). This combination of measurement approaches increases confidence in our results by providing what Campbell referred to as “convergent validity” (Campbell and Fiske, 1959).

We found that our police-officer participants’ driving performance was significantly impacted by fatigue. Both the objective PVT and the subjective KSS measures predicted changes in the probability of a collision in the study – with PVT producing a stronger prediction than KSS. This indicates that these relatively simple tools could be used in field settings to assess officers’ fatigue levels. However, stopping to take a PVT or fill out the KSS often may be impractical in the field. For this reason, well-established performance measures taken while driving, such as lane deviation and braking latency, could also be used to predict collisions. These measures are important leading indicators of decay in driving performance that could provide a basis for in-vehicle early warning systems. These early warning systems could be used to prevent collisions by monitoring performance changes before an officer is too impaired to do more than pull off the road.

Braking latency also can be obtained from the data stream on the patrol vehicle’s on-board computer. Although lane deviation is difficult to measure consistently in operational settings, it can be approximated very well by measuring steering-wheel movement variability (Forsman et al., 2013) – which can also be obtained from the on-board computer.

The shifts that officers worked also significantly affected driving performance and the probability of a collision. Officers working night shifts (swing and graves) had more collisions and greater lane deviation than those working day shifts (days and power). Night shift officers also took longer to react to the PVT stimuli and had more lapses of attention.

These findings are important because although most policing agencies have policies, practices, and training intended to improve workplace driving safety, drowsy driving on the way home from work has largely been overlooked. Our findings indicate that this may be a critical oversight and that drowsy police officers – especially those working night shifts – are at greater risk for collisions.

Policing agencies should develop policies to help ensure that officers are not too tired to drive home safely after shift. Practical measures employed in many operational settings include: establishment of a post-shift napping room, providing a ride home to officers who are likely to be overly drowsy, training officers to manage the challenges associated with shift work and fatigue, and establishing supervisor and peer-based monitoring for post-shift fatigue risks (Table IX).
Experimental process used during fatigued and rested conditions

<table>
<thead>
<tr>
<th>Activity</th>
<th>Fatigued condition study day (11.8 hr)</th>
<th>Rested condition study day (5.3 hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time line (min)</td>
<td>60 12 16 12 16 12 45 12 60 12 30 12 20</td>
<td>60 12 16 12 16 12 45 12 60 12 30 12 20</td>
</tr>
<tr>
<td>Activity</td>
<td>Intake PVT Patrol drive simulation</td>
<td>Intake PVT Patrol drive simulation</td>
</tr>
<tr>
<td></td>
<td>PVT Patrol drive simulation</td>
<td>PVT Patrol drive simulation</td>
</tr>
<tr>
<td></td>
<td>PVT Deadly force simulation</td>
<td>PVT Cognitive battery</td>
</tr>
<tr>
<td></td>
<td>PVT Drive home simulation</td>
<td>PVT Drive home simulation</td>
</tr>
<tr>
<td></td>
<td>PVT Meal</td>
<td>PVT Meal</td>
</tr>
<tr>
<td></td>
<td>Sleep in lab</td>
<td>Sleep in lab</td>
</tr>
<tr>
<td></td>
<td>Debrief</td>
<td>Debrief</td>
</tr>
</tbody>
</table>

**Note:** During the fatigued condition, participants reported to the laboratory immediately after their work shift, approximately 20 minutes after coming off duty. Once they had been equipped with measurement devices, they completed the series of experimental tasks. Afterward, they slept in dark, sound-proof, comfortable, and private rooms to assure that they were sufficiently rested to drive home safely. In the rested condition, they reported to the lab at the same time of day and completed the same tasks in the same order, but were allowed to go home immediately after the study. Experimental activities associated with the research reported here are shaded in gray.
Notes
1. “Graves” or the “graveyard” shift typically starts during the early evening and extends through to the next morning.
2. This agency’s patrol division does not respond to property crime due to this understaffing.
3. Approved by the Washington State University Institutional Review Board, and funded by the California Commission on Peace Officer Standards and Training (contract no. 00112338) and the Office of Naval Research (grant no. N000141110185).
4. The emergency braking instances in these scenarios were designed to allow a vigilant driver enough time and distance to react to the stimulus and apply their brakes for sufficient time to slow or stop the vehicle before colliding with the pedestrian or dog. This was possible at speeds up to 64 mph thus allowing for momentary fluctuations in speed.

References
Bureau of Justice Statistics (BJS) (2013), Local Police Departments, Department of Justice, NCJ 231174, Washington, DC.


National Transportation Safety Board (NTSB) (1990), *Safety Study: Fatigue, Alcohol, Other Drugs, and Medical Factors in Fatal-to-the-Driver Heavy Truck Crashes*, Vol. 1, National Transportation Safety Board, Washington, DC.


Further reading


About the authors

Stephen M. James is a Doctoral candidate in the Department of Criminal Justice and Criminology at the Washington State University. He has managed the Simulated Hazardous Operational Tasks Laboratory at the Sleep and Performance Center, Washington State University, Spokane. James has been a member of the CA POST Safe Driving Research Team since 2009. Stephen M. James is the corresponding author and can be contacted at: stevejames@wsu.edu
Bryan Vila PhD, is a Professor of Criminal Justice and Criminology at the WSU Spokane. Prior to joining WSU in July 2005, he directed the Division of Crime Control and Prevention Research at the US Department of Justice’s National Institute of Justice. Dr Vila pioneered the study of police fatigue. He has served as a Principal Investigator on 16 externally funded research projects and published four books and more than 60 refereed articles and book chapters. Dr Vila received his PhD in Ecology from the University of California, Davis in 1990. Before he became an academic, he served as a Law Enforcement Officer for 17 years in local, national, and international settings.
This article has been cited by:

1. Bryan Vila, Charles Samuels, Nancy J. Wesensten. Sleep Problems in First Responders and in Deployed Military Personnel 726-735.e4. [CrossRef]